Realistic Image Synthesis

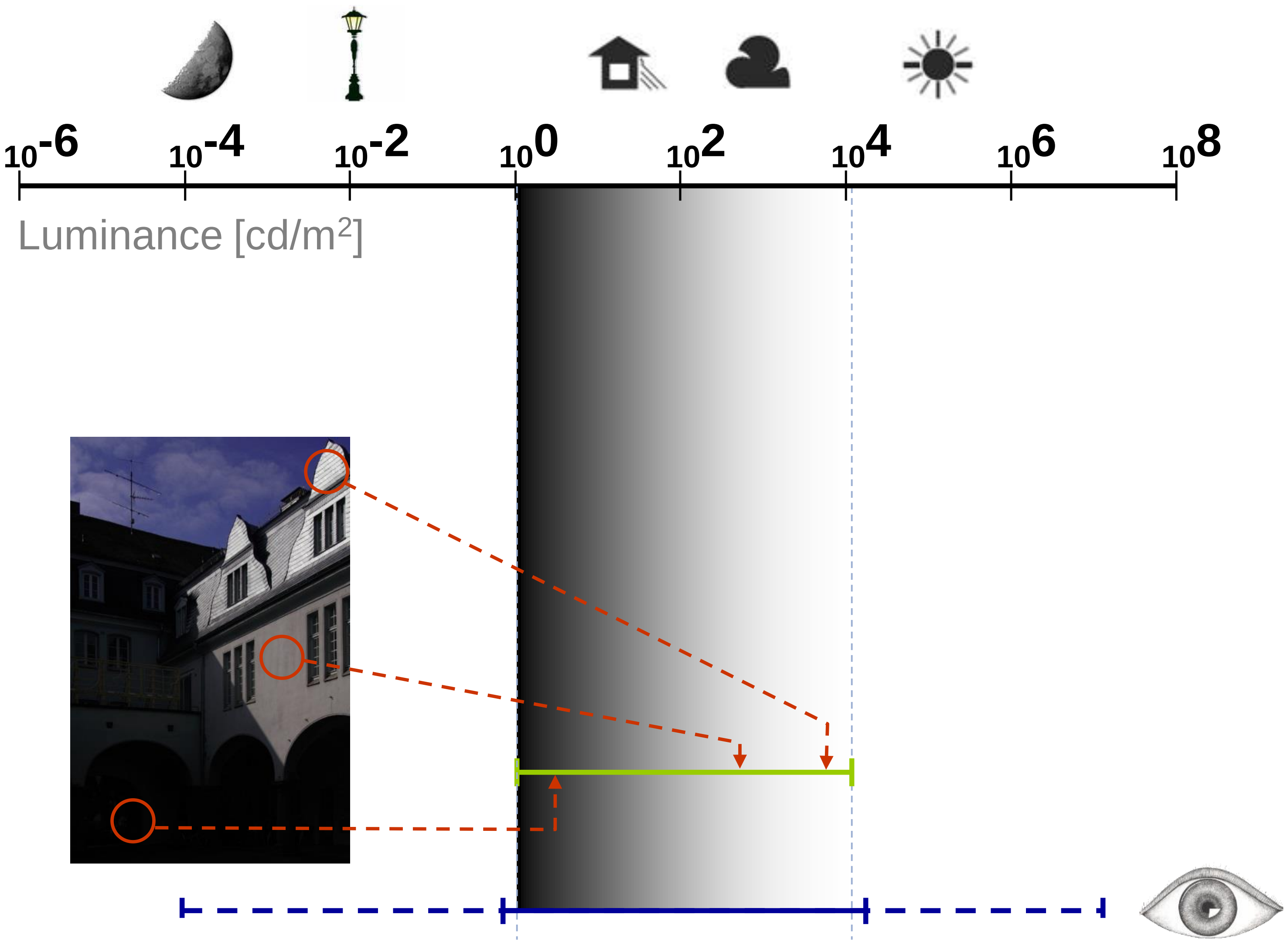
- HDR Capture & Tone Mapping -

Philipp Slusallek
Karol Myszkowski
Gurprit Singh

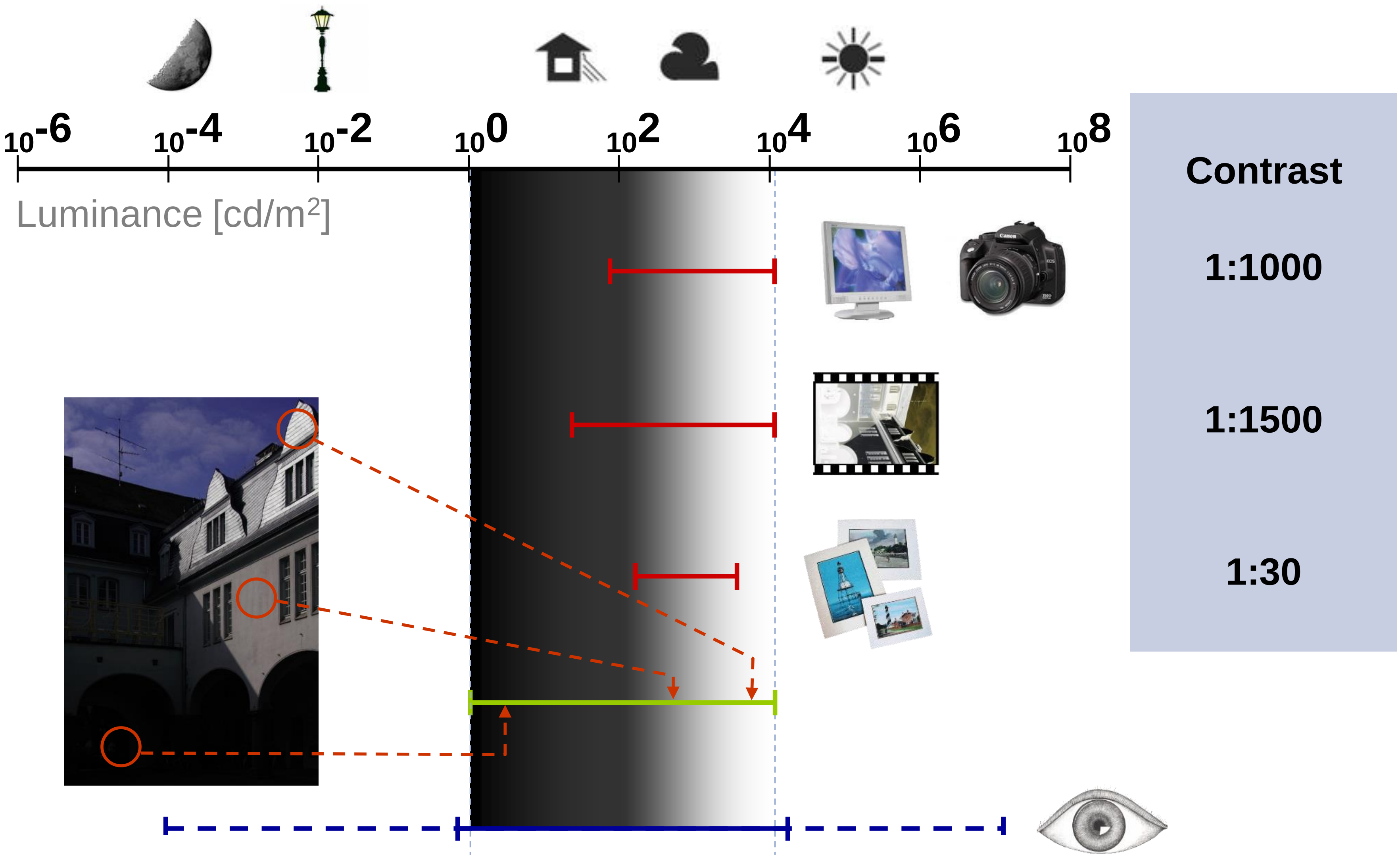
LDR vs HDR – Comparison

Standard Dynamic Range		High Dynamic Range
	QUALITY OF CONTRAST & COLOR	
50 dB	Camera Dynamic Range	120 dB
1:200	Display Contrast	1:15.000
limited	Color Gamut	vivid and saturated colors
display-referred	Image Representation	scene-referred
display limited	Fidelity	as good as the eye can see

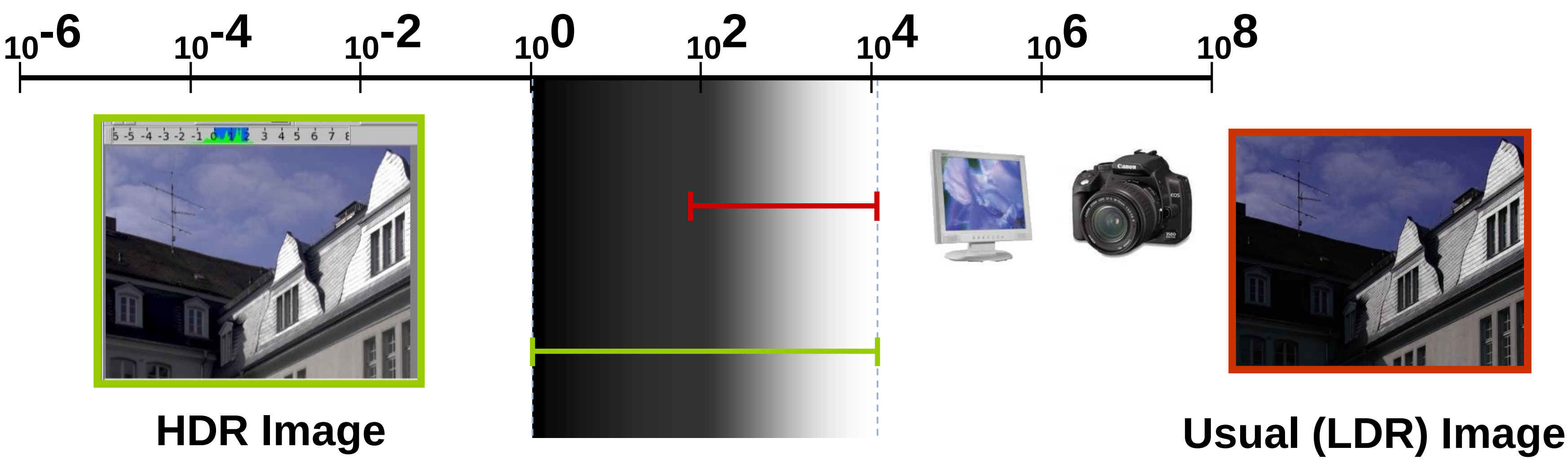
Various Dynamic Ranges (1)



Various Dynamic Ranges (2)



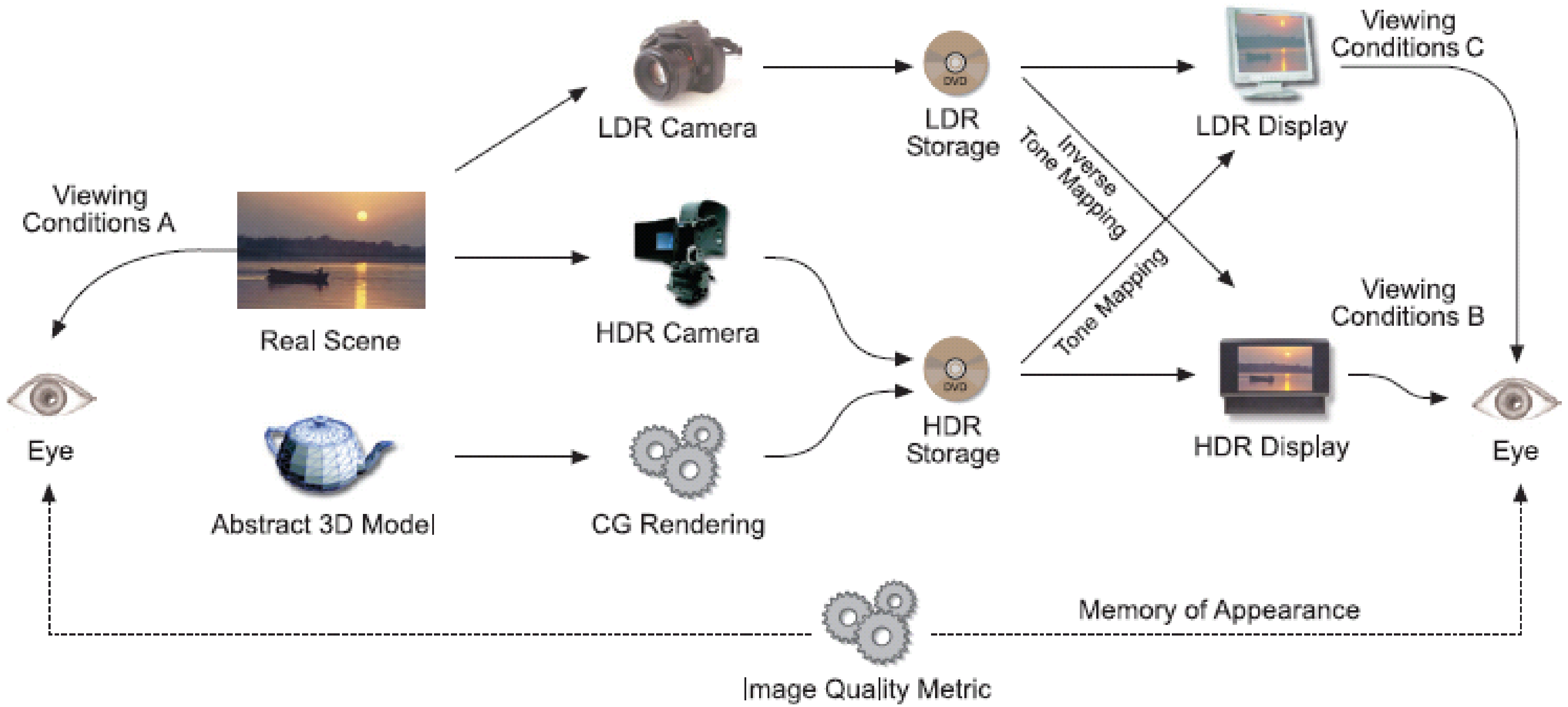
High Dynamic Range



Measures of Dynamic Range

Contrast ratio	$\mathbf{CR = 1 : (Y_{peak}/Y_{noise})}$	displays (1:500)
Orders of magnitude	$\mathbf{M = \log_{10}(Y_{peak})-\log_{10}(Y_{noise})}$	HDR imaging (2.7 orders)
Exposure latitude (f-stops)	$\mathbf{L = \log_2(Y_{peak})-\log_2(Y_{noise})}$	photography (9 f-stops)
Signal to noise ratio (SNR)	$\mathbf{SNR = 20*\log_{10}(A_{peak}/A_{noise})}$	digital cameras (53 [dB])

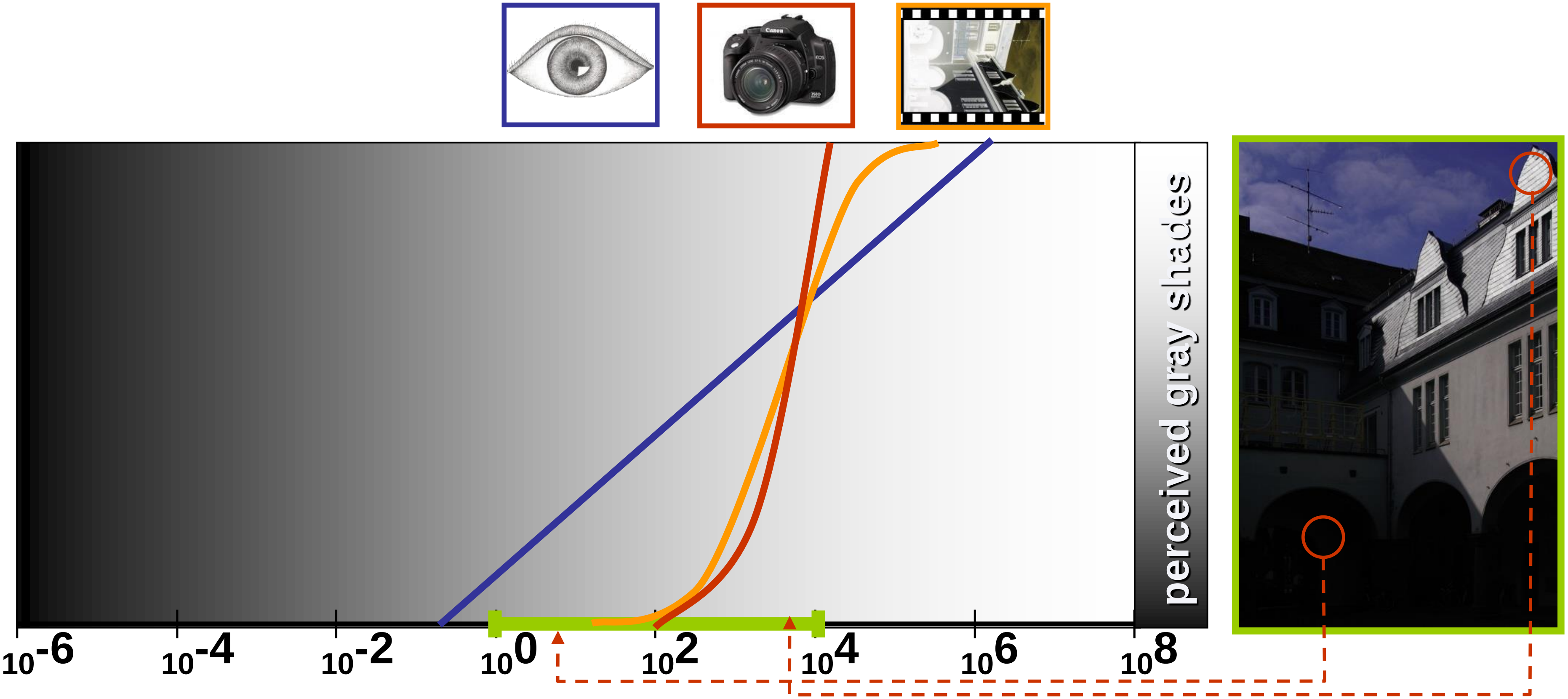
HDR Pipeline



Lecture Overview

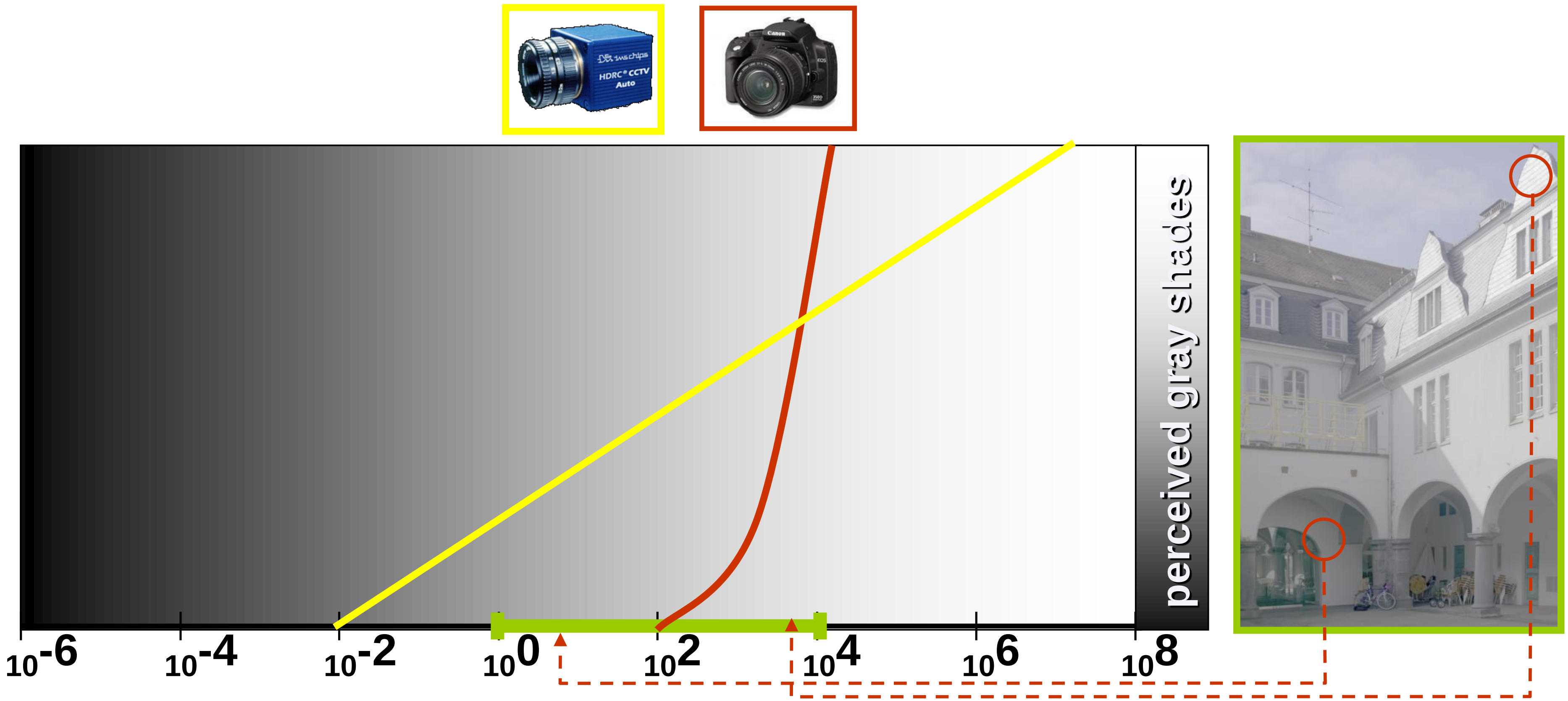
- Capture of HDR images and video
  - Multi-exposure techniques
  - Photometric calibration
- Tone Mapping of HDR images and video
  - Early ideas for reducing contrast range
  - Image processing – fixing problems
  - Alternative approaches
  - Perceptual effects in tone mapping
- Summary

HDR: a normal camera can't…



- linearity of the CCD sensor
- bound to 8-14bit processors
- saved in an 8bit gamma corrected image

HDR Sensors



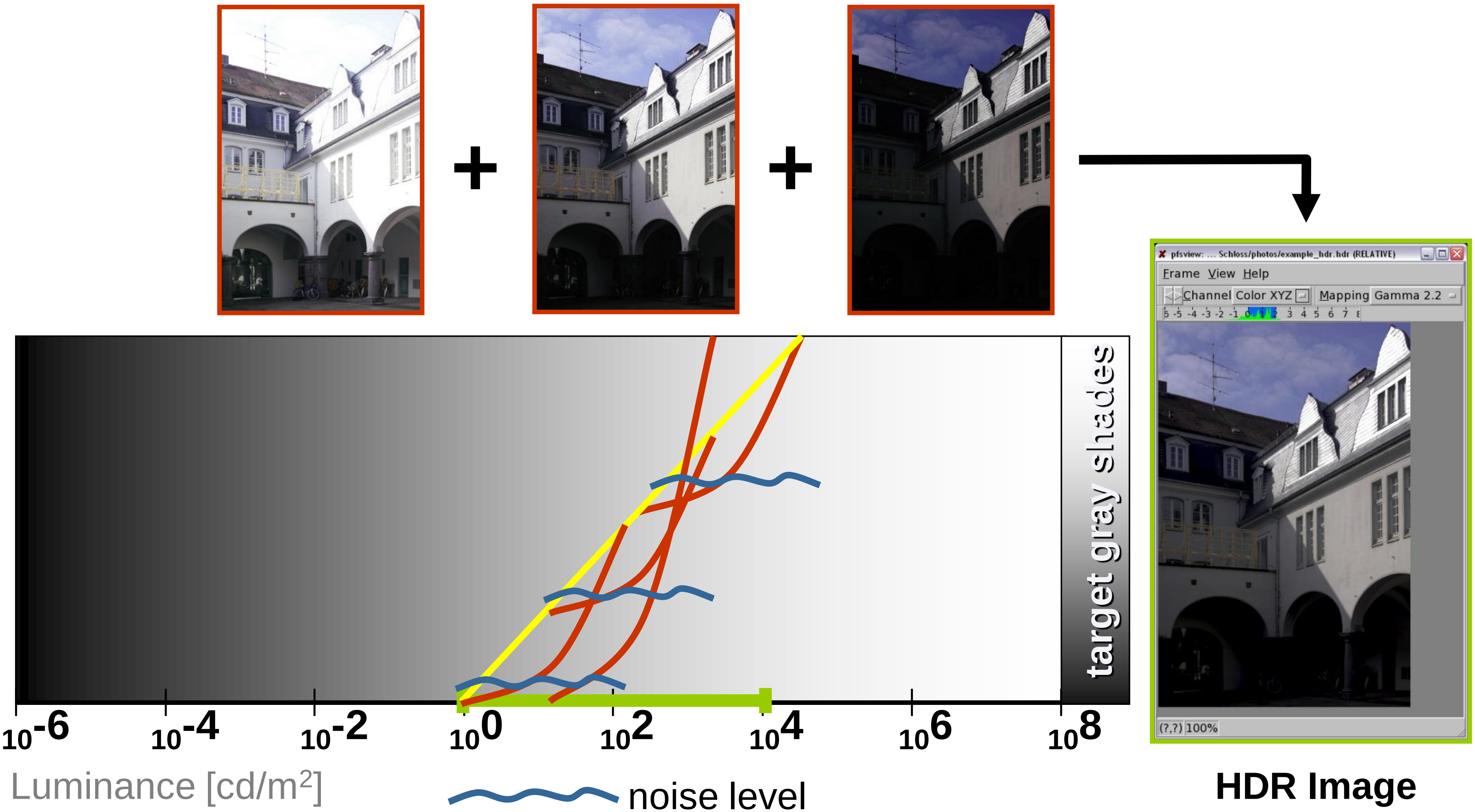
- logarithmic response
- locally auto-adaptive
- hybrid sensors (linear-logarithmic)

HDR with a normal camera

Dynamic range of a typical CCD	**1:1000**
Exposure variation ($^1/60$: $^1/6000$)	1:100
Aperture variation (f/2.0 : f/22.0)	~1:100
Sensitivity variation (ISO 50 : 800)	~1:10
Total operational range	**1:100,000,000** **High Dynamic Range!**

Dynamic range of a single capture only **1:1000.**

Multi-exposure Technique (1)



Multi-exposure Technique (2)

- Input
  - images captured with varying exposure
    - change exposure time, sensitivity (ISO), ND filters
    - same aperture!
    - exactly the same scene!
- **Unknowns**
  - camera response curve (can be given as input)
  - HDR image
- Process
  - recovery of camera response curve (if not given as input)
  - linearization of input images (to account for camera response)
  - normalization by exposure level
  - suppression of noise
  - estimation of HDR image (linear combination of input images)

Algorithm (1/3)

Camera Response

$$y_{ij} = I(x_{ij} \cdot t_i)$$

Merge to HDR

- Linearize input images and normalize by exposure time

$$x_{ij} = \frac{I^{-1}(y_{ij})}{t_i}$$

assume $\boldsymbol{I}$ is correct (initial guess)

- Weighted average of images (weights from certainty model)

$$x_j = \frac{\sum_i w_{ij} x_{ij}}{\sum_i w_{ij}}$$

Optimize Camera Response

- Camera response

$$I^{-1}(y_{ij}) = t_i x_j$$

assume $\boldsymbol{x_j}$ is correct

- Refine initial guess on response
  - linear eq. (Gauss-Seidel method)

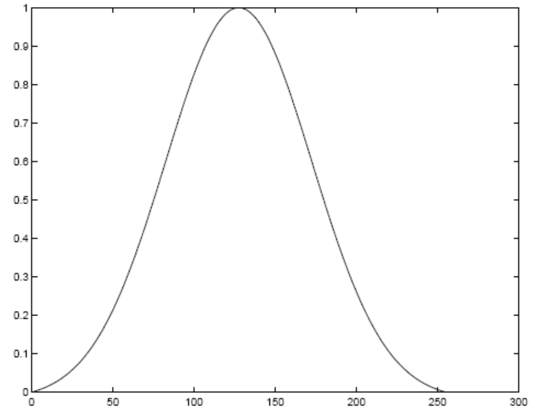
$$E_m = \{(i, j) : y_{ij} = m\}$$

$$I^{-1}(m) = \frac{1}{\mathrm{Card}(E_m)} \sum_{i, j \in E_m} t_i x_j$$

$\boldsymbol{t_i}$ exposure time of image $\boldsymbol{i}$
$\boldsymbol{y_{ij}}$ pixel of input image $\boldsymbol{i}$ at position $\boldsymbol{j}$
$\boldsymbol{I}$ camera response
$\boldsymbol{x_j}$ HDR image at position $\boldsymbol{j}$
$\boldsymbol{w}$ weight from certainty model
$\boldsymbol{m}$ camera output value

Algorithm (2/3)

- Certainty model (for 8bit image)
  - High confidence in middle output range
  - Dequantization uncertainty term
  - Noise level



$$w(y_{ij}) = \exp\left(-4\frac{(y_{ij} - 127.5)^2}{127.5^2}\right)$$

- Longer exposures are favored t_i^2
  - Less random noise
- Weights

$$w_{ij} = w(y_{ij})t_i^2$$

Algorithm (3/3)

1. Assume initial camera response $\boldsymbol{I}$ (linear)
2. Merge input images to HDR

$$x_j = \frac{\sum_i \boxed{w(y_{ij})t_i^2} \cdot \boxed{\frac{I^{-1}(y_{ij})}{t_i}}}{\sum_i w(y_{ij})t_i^2}$$

3. Refine camera response

$$E_m = \{(i,j) : y_{ij} = m\}$$

$$I^{-1}(m) = \frac{1}{\mathrm{Card}(E_m)} \sum_{i,j \in E_m} t_i x_j$$

4. Normalize camera response by middle value: $I^{-1}(m)/I^{-1}(m_{med})$
5. Repeat 2,3,4 until objective function is acceptable

$$O = \sum_{i,j} w(y_{ij})(I^{-1}(y_{ij}) - t_i x_j)^2$$

Other Algorithms

- [Debevec & Malik 1997]
  - in log space
  - assumptions on the camera response
    - monotonic
    - continuous
  - a lot to compute for >8bit
- [Mitsunaga & Nayar 1999]
  - camera response approximated with a polynomial
  - very fast
- Both are more robust but less general
  - not possible to calibrate non-standard sensors

Issues with Multi-exposures

- How many source images?
  - First expose for shadows: all output values above 128 (for 8bit imager)
  - 2 f-stops spacing (factor of 4) between images
  - one or two images with 1/3 f-stop increase will improve quantization in HDR image
  - Last exposure: no pixel in image with maximum value
- Alignment
  - Shoot from tripod
  - Otherwise use panorama stitching techniques to align images
- Ghosting
  - Moving objects between exposures leave “ghosts”
  - Statistical method to prevent such artifacts
- Practical only for images!
  - Multi-exposure video projects exist, but require care with subsequent frame registration by means of optical flow

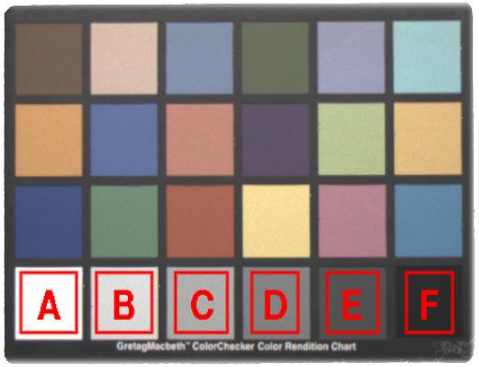
Photometric Calibration

- Converts camera output to luminance
  - requires camera response,
  - and a reference measurement for known exposure settings
- Applications
  - predictive rendering
  - simulation of human vision response to light
  - common output in systems combining different cameras

Calibration (Response Recovery)

- Camera response can be reused
  - for the same camera
  - for the same picture style settings (eg. contrast)
- Good calibration target
  - Neutral target (e.g. Gray Card)
    - Minimize impact of color processing in camera
  - Smooth illumination
    - Uniform histogram of input values
  - Out-of-focus
    - No interference with edge aliasing and sharpening

Photometric Calibration (cntd.)

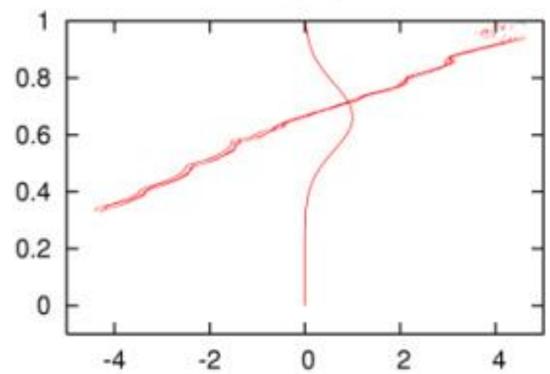


acquire target

camera output values

measure luminance

luminance values



camera response

HDR Sensor vs. Multi-exposure

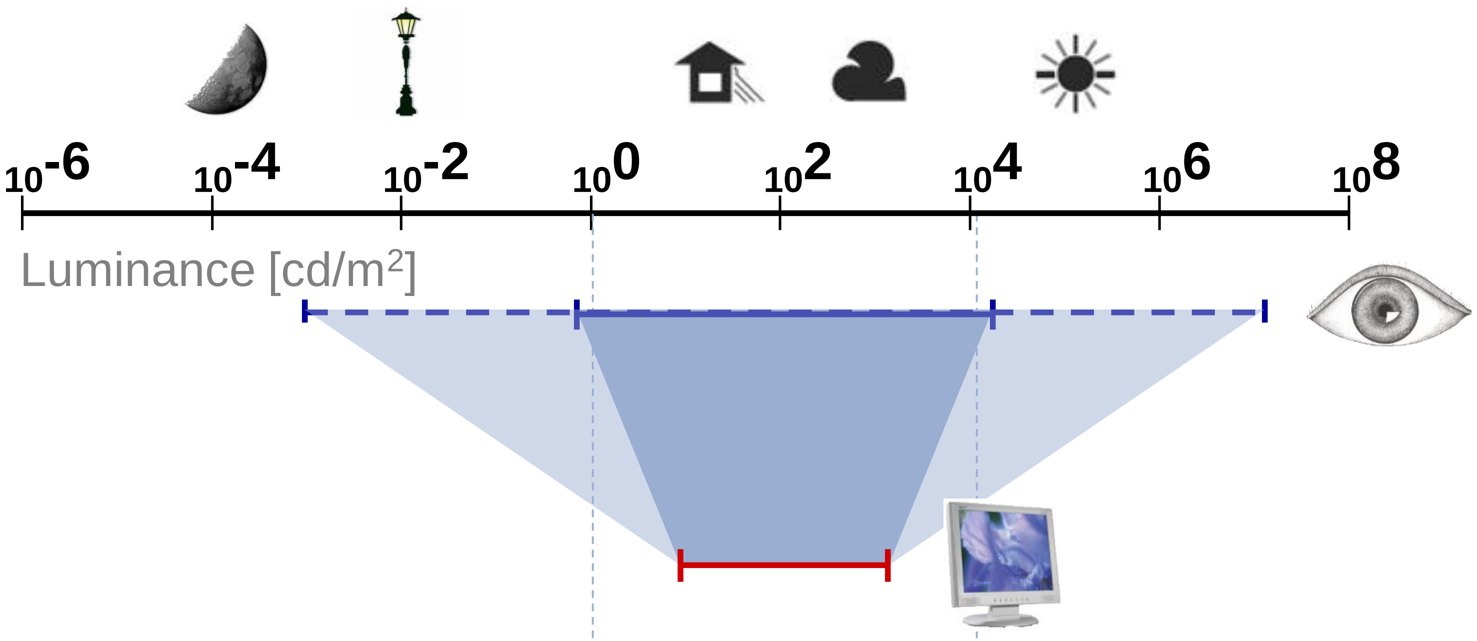
- HDR camera
  - Fast acquisition of dynamic scenes at 25fps without motion artifacts
  - Currently lower resolution

- LDR camera + multi-exposure technique
  - Slow acquisition (impossible in some conditions)
  - Higher quality and resolution
  - High accuracy of measurements

Lecture Overview

- Capture of HDR images and video
  - HDR sensors
  - Multi-exposure techniques
  - Photometric calibration
- Tone Mapping of HDR images and video
  - Early ideas for reducing contrast range
  - Image processing – fixing problems
  - Alternative approaches
  - Perceptual effects in tone mapping
- Summary

HDR Tone Mapping



- Objectives of tone mapping
  - nice looking images
  - perceptual brightness match
  - good detail visibility
  - equivalent object detection performance
  - really application dependent…

General Idea

- Luminance as an input
  - absolute luminance
  - relative luminance (luminance factor)
- Transfer function
  - maps luminance to a certain pixel intensity
  - may be the same for all pixels (**global operators**)
  - may depend on spatially local neighbors (**local operators**)
  - dynamic range is reduced to a specified range
- Pixel intensity as output
  - often requires gamma correction
- Colors
  - most algorithms work on luminance
    - use RGB to Yxy color space transform
    - inverse transform using tone mapped luminance
  - otherwise each RGB channel processed independently

General Problems

- Constraints in observation conditions
  - limited contrast
  - quantization
  - different ambient illumination
  - different luminance levels
  - adaptation level often incorrect for the scene
  - narrow field of view
- Appearance may not always be matched

Transfer Functions

- Linear mapping (naïve approach)
  - like taking a usual photo
- Brightness function
- Sigmoid responses
  - simulate our photoreceptors
  - simulate response of photographic film
- Histogram equalization
  - standard image processing
  - requires detection threshold limit to prevent contouring

Adapting Luminance

- Maps luminance on a scale of gray shades
- Task is to match gray levels
  - average luminance in the scene is perceived as a gray shade of medium brightness
  - such luminance is mapped on medium brightness of a display
  - the rest is mapped proportionally
- Practically adjusts brightness
  - sort of like using gray card or auto-exposure in photography
  - goal of adaptation processes in human vision
- Adapting luminance exists in many TM algorithms

$$Y_A = \exp\left(\frac{\sum \log(Y + \varepsilon)}{N} - \varepsilon \right)$$

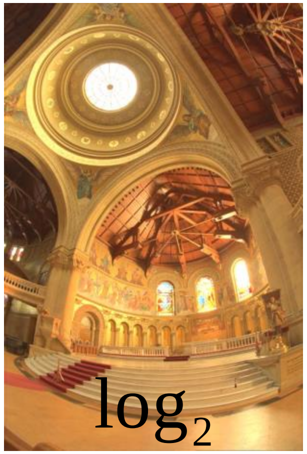
Logarithmic Tone Mapping

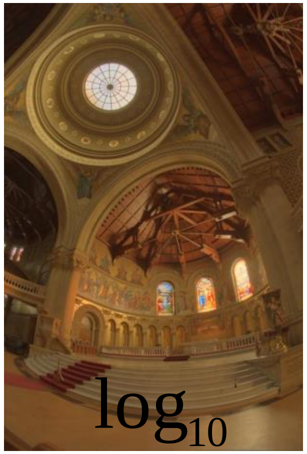
- Logarithm is a crude approximation of brightness
- Change of base for varied contrast mapping in bright and dark areas
  - $\log_{10}$ maps better for bright areas
  - $\log_2$ maps better for dark areas
- Mapping parameter *bias* in range 0.1:1

$$Y' = \frac{Y}{Y_A}$$

$$L = L_{\max} \cdot \frac{\log_{base(Y)}\left(Y'+1\right)}{\log_{10}\left(\max(Y')+1\right)}$$

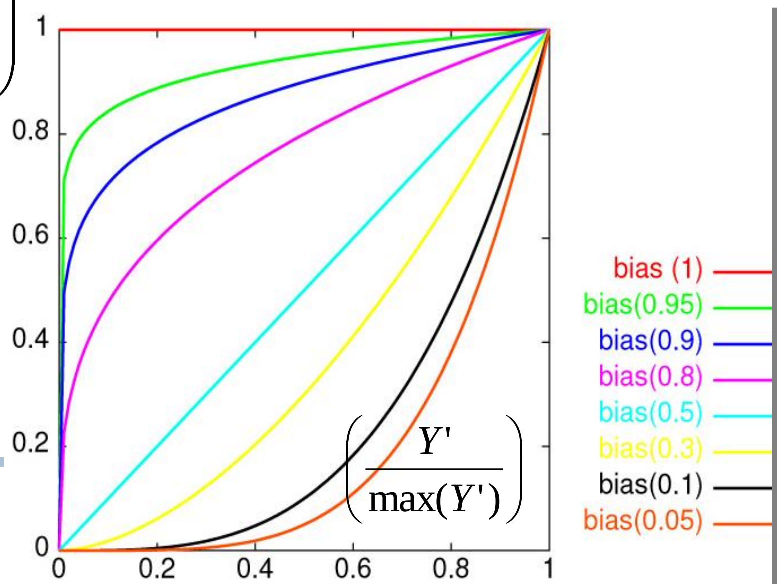
$$base(Y') = 2 + 8 \cdot \left(\frac{Y'}{\max(Y')}\right)^{\log_{0.5} bias}$$





Logarithmic Tone Mapping

$$\left(\frac{Y'}{\max(Y')}\right)^{\log_{0.5} bias}$$



- These images illustrate how high luminance values are clamped to the maximum displayable values using different bias parameter values.
- The scene dynamic range is 1:11,751,307.

Sigmoid Response

- Model of photoreceptor

$$L = \frac{Y}{Y + (f \cdot Y_A)^m} L_{\max}$$

- Brightness parameter $\boldsymbol{f}$
- Contrast parameter $\boldsymbol{m}$
- Adapting luminance $\boldsymbol{Y_A}$
  - average in an image
  - measured pixel (equal to $\boldsymbol{Y}$)

logarithmic mapping

sigmoid mapping

Histogram Equalization (1)

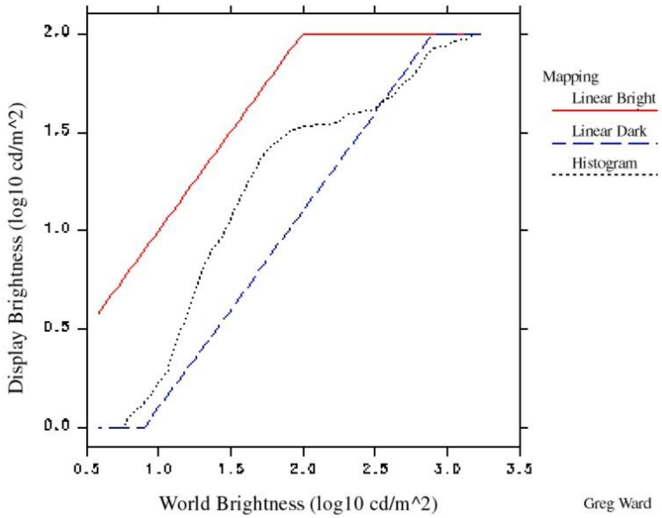
- Adapts transfer function to distribution of luminance in the image
- Algorithm:
  - compute histogram
  - compute transfer function (cumulative distribution)
  - limit slope of transfer function to prevent contouring
    - contouring – visible difference between 1 quantization step
    - use threshold versus intensity function (TVI)

      TVI gives visible luminance difference for adapting luminance
- Most optimal transfer function
- Not efficient when large uniform areas are present in the image

Histogram Equalization (2)

A linear mapping of the luminances that overexposes the view through the window.

World to Display Luminance Mapping



A linear mapping of the luminances that underexposes the view of the interior.

The luminances mapped to preserve the visibility of both indoor and outdoor features.

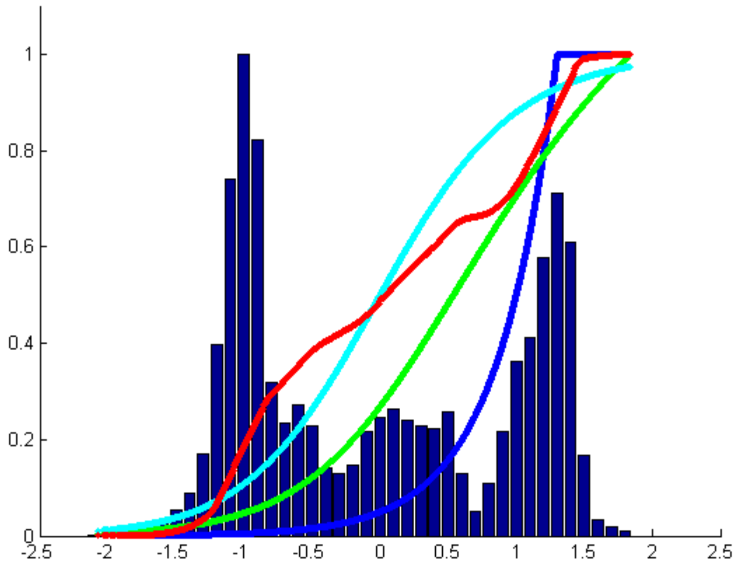
Transfer Functions Compared


linear


logarithmic


photoreceptor


histogram eq.



- Interpretation
  - steepness of slope is contrast
  - luminance for which output is ~0 and ~1 is not transferred
- Usually low contrast for dark and bright areas!

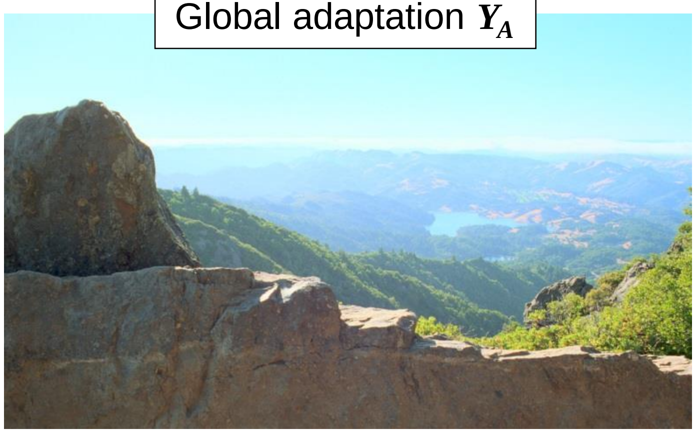
Problem with Details

- Strong compression of contrast puts micro-contrasts (details) below quantization level

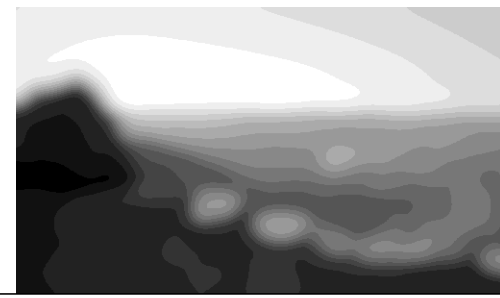
Introducing Local Adaptation

- Eye adapts locally to observed area

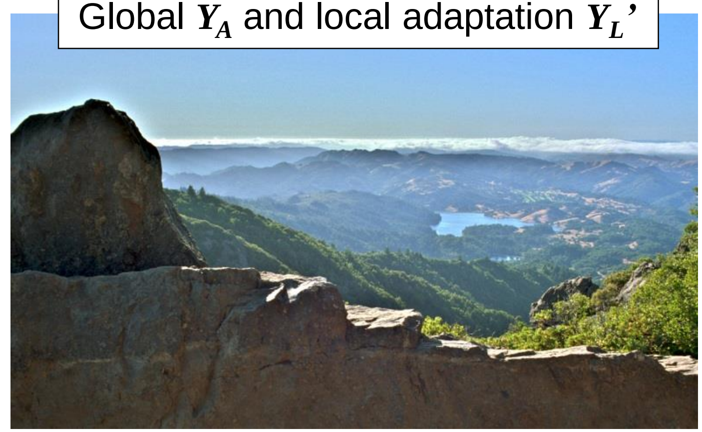
$$L = \frac{Y'}{Y'+1} \longleftarrow \quad Y' = \frac{Y}{Y_A} \quad \longrightarrow L = \frac{Y'}{Y_L'+1}$$



Global adaptation Y_A

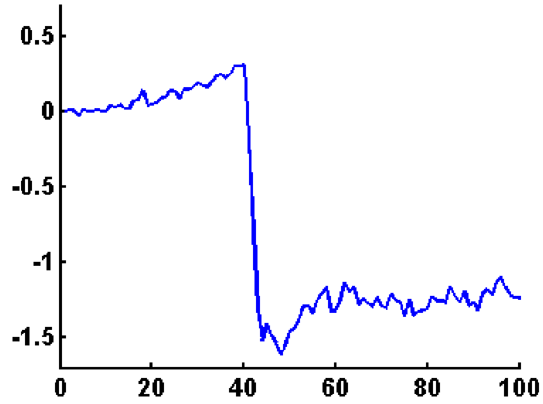


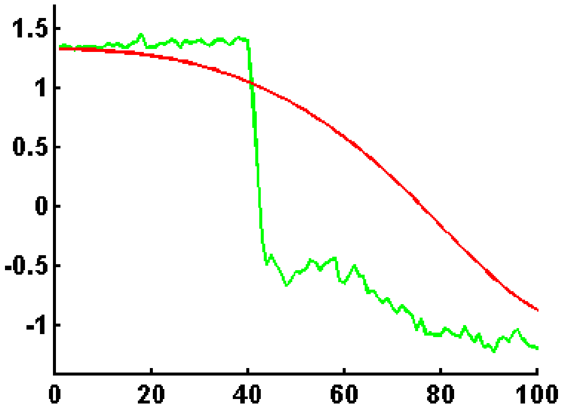
Gaussian blur of HDR image, σ ~ 1deg of visual angle.



Global Y_A and local adaptation Y_L’

The Halo Artifact



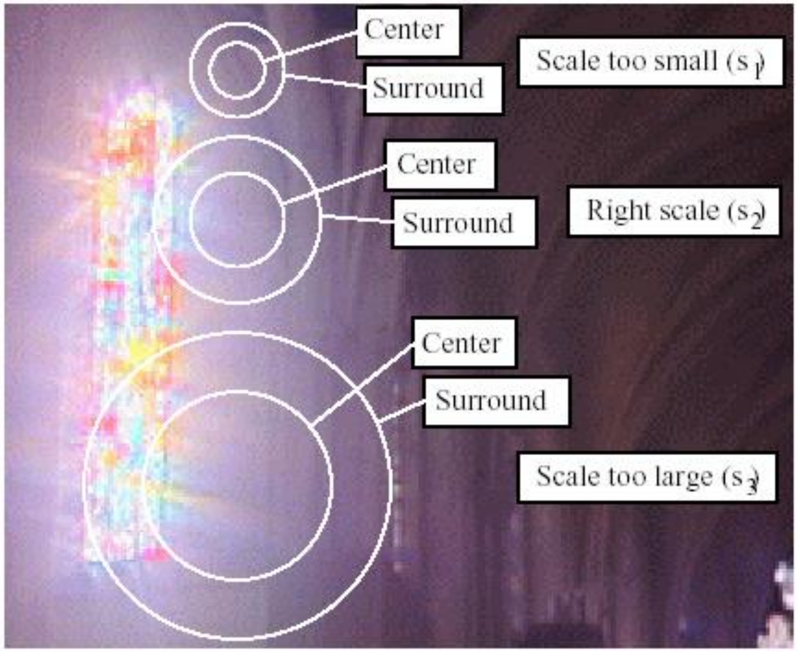


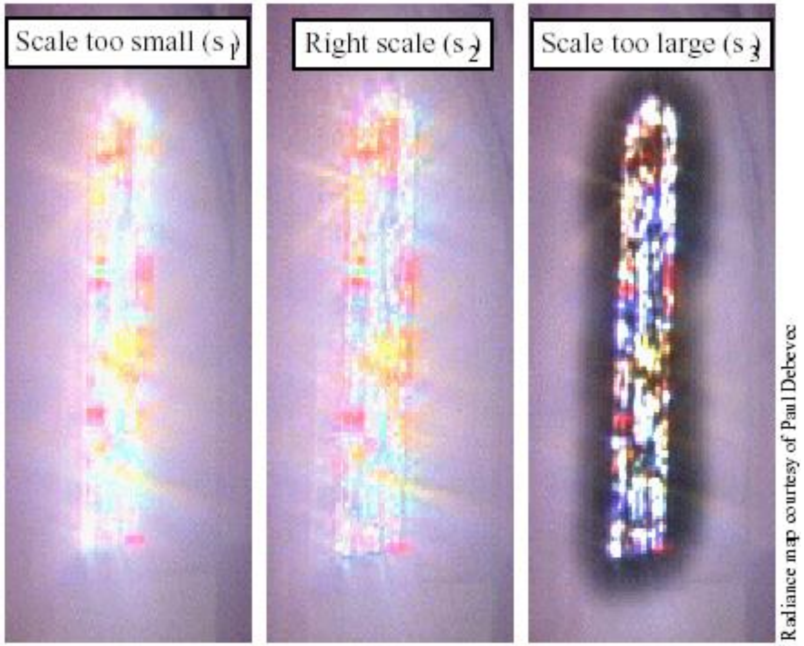
- Scan line example:
  - Gaussian blur under- (over-) estimates local adaptation near a high contrast edge
  - tone mapped image gets too bright (too dark) closer to such an edge
- Smaller blur kernel reduces the artifact (but then no details)
- Larger blur kernel spreads the artifact on larger area

Adjusting Gaussian Blur

- So called: Automatic Dodging and Burning
  - for each pixel, test increasing blur size $\boldsymbol{\sigma_i}$
  - choose the largest blur which does not show halo artifact

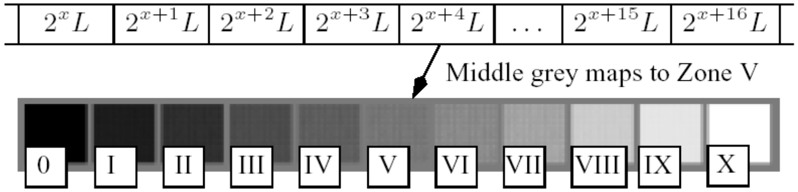
$$\left| Y_L(x,y,\sigma_i) - Y_L(x,y,\sigma_{i+1}) \right| < \varepsilon$$





Photographic Tone Reproduction

- Map luminance using Zone System



Print zones: Zone V 18% reflectance

$$Y' = \frac{Y}{Y_A}, \quad Y_A = \exp\left(\frac{\sum \log(Y)}{N} \right)$$

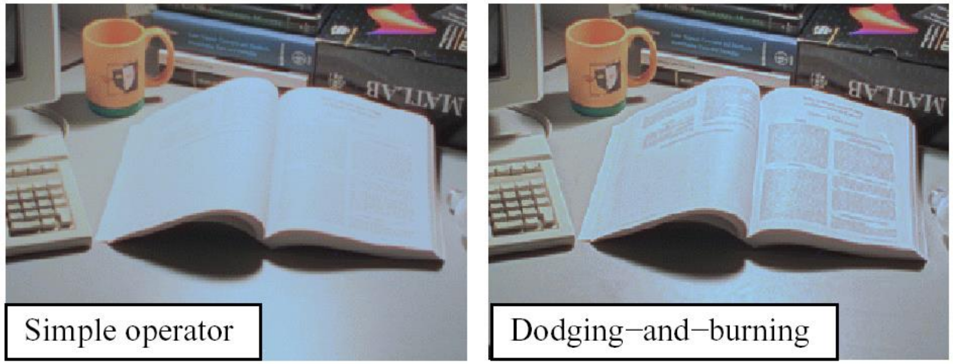
- Find local adaptation for each pixel
  - appropriate size of Gaussian (automatic dodging & burning)

$$\left| Y_L'(x,y,\sigma_i) - Y_L'(x,y,\sigma_{i+1}) \right| < \varepsilon$$

- Tone map using sigmoid function
  - different blur levels from Gaussian pyramid

$$L(x,y) = \frac{Y'(x,y)}{Y_L'(x,y,\sigma_{x,y}) + 1}$$

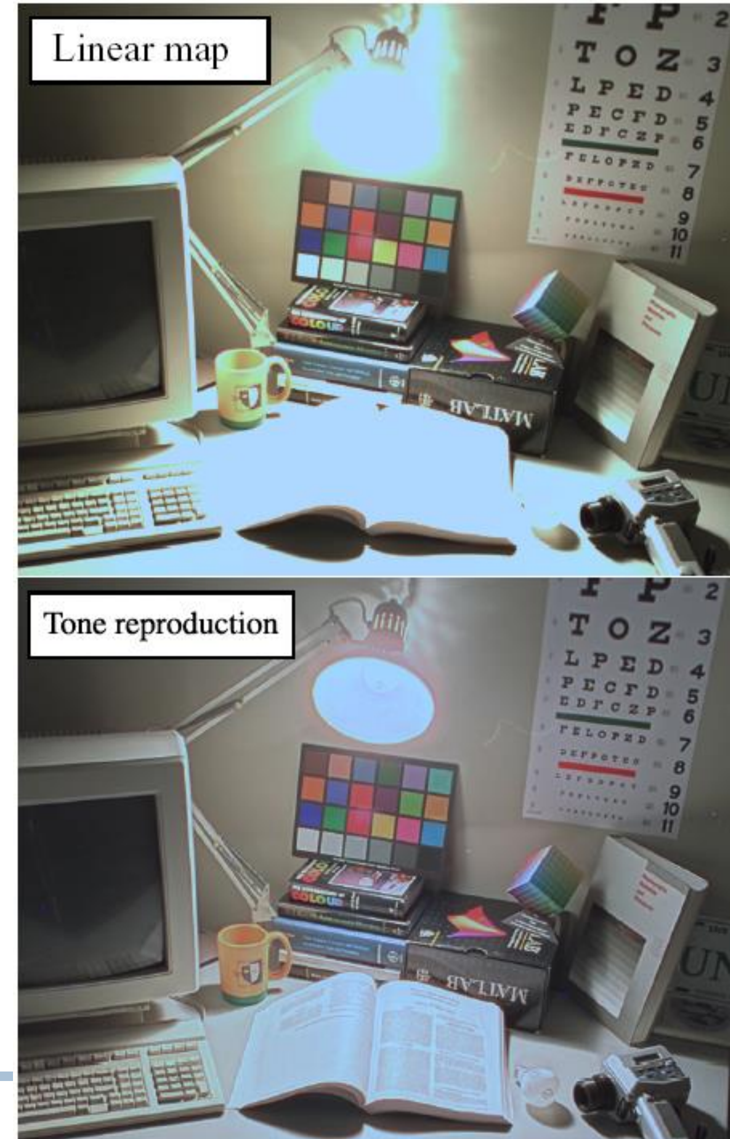
Photographic Tone Reproduction



dodge luminance of pixels in bright regions is significantly decreased

burn pixels in dark regions are compressed less, so their relative intensity increases

Automatic dodging-and-burning technique is more effective in preserving local details (notice the print in the book).

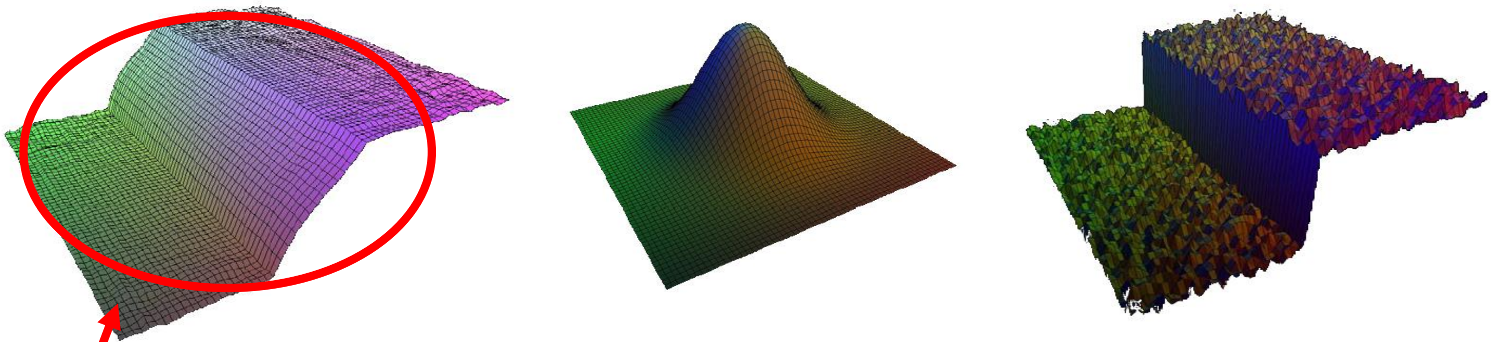


Bilateral Filtering

- Edge preserving Gaussian filter to prevent halo
- Conceptually based on intrinsic image models:
  - decoupling of illumination and reflectance layers
    - very simple task in CG
    - complicated for real-world scenes
  - compress range of illumination layer
  - preserve reflectance layer (details)
- Bilateral filter separates:
  - texture details (high frequencies, low amplitudes)
  - illumination (low frequencies, high contrast edges)

Illumination Layer (1)

- Identify low frequencies in the scene
  - Gaussian filtering leads to halo artifacts



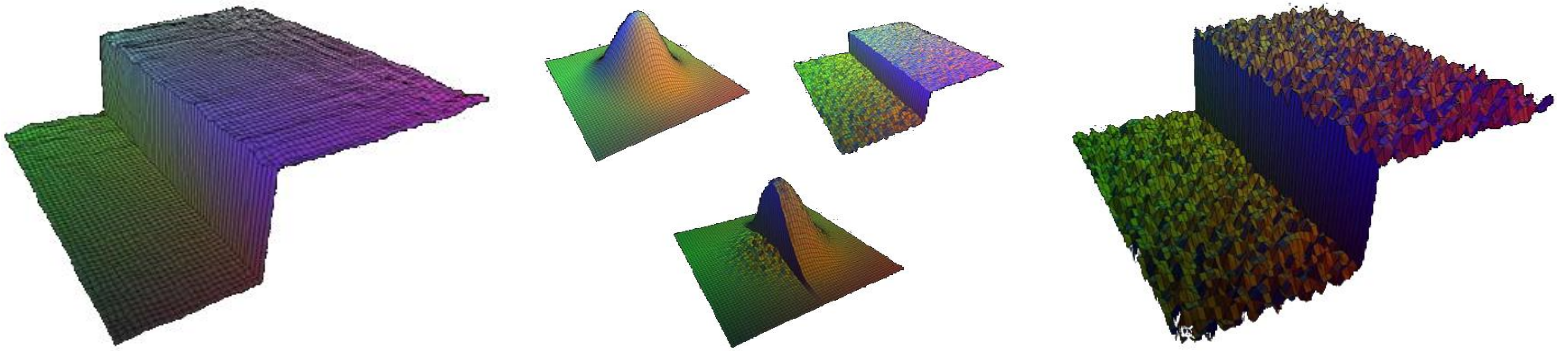
$$J_p = \frac{1}{W_p} \sum_{q \in N(p)} f_{\sigma_s}\left(\|p-q\|\right) \cdot I_q$$

f **spatial kernel with large** σ_s

lost sharp edge

Illumination Layer (2)

- Edge preserving filter – no halo artifacts

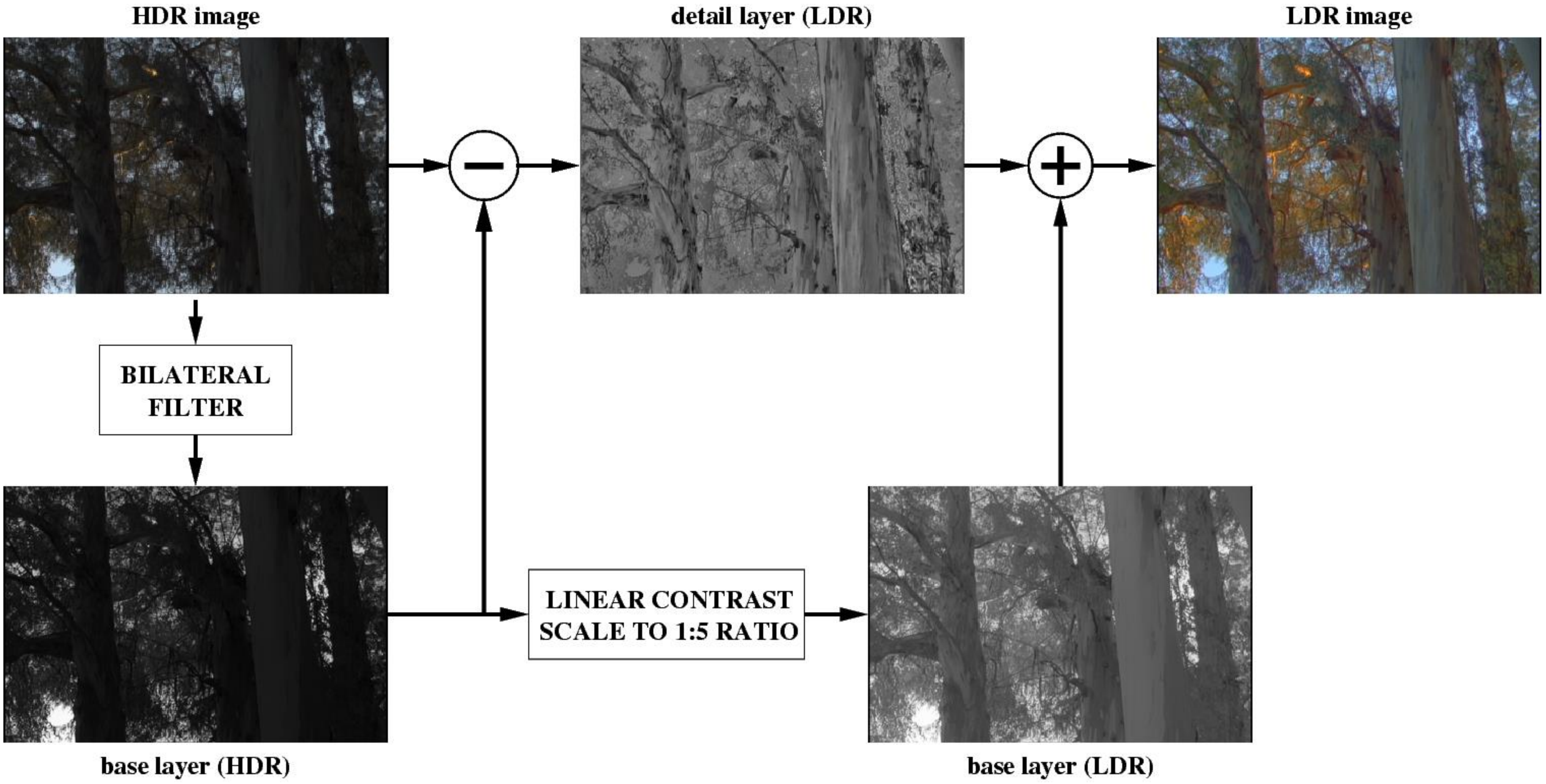


$$J_p = \frac{1}{W_p} \sum_{q \in N(p)} f_{\sigma_s}\left(\left\| p - q \right\|\right) \cdot g_{\sigma_r}\left(\left| I_p - I_q \right|\right) \cdot I_q$$

f **spatial kernel with large σ_s**

g **range kernel with very small σ_r**

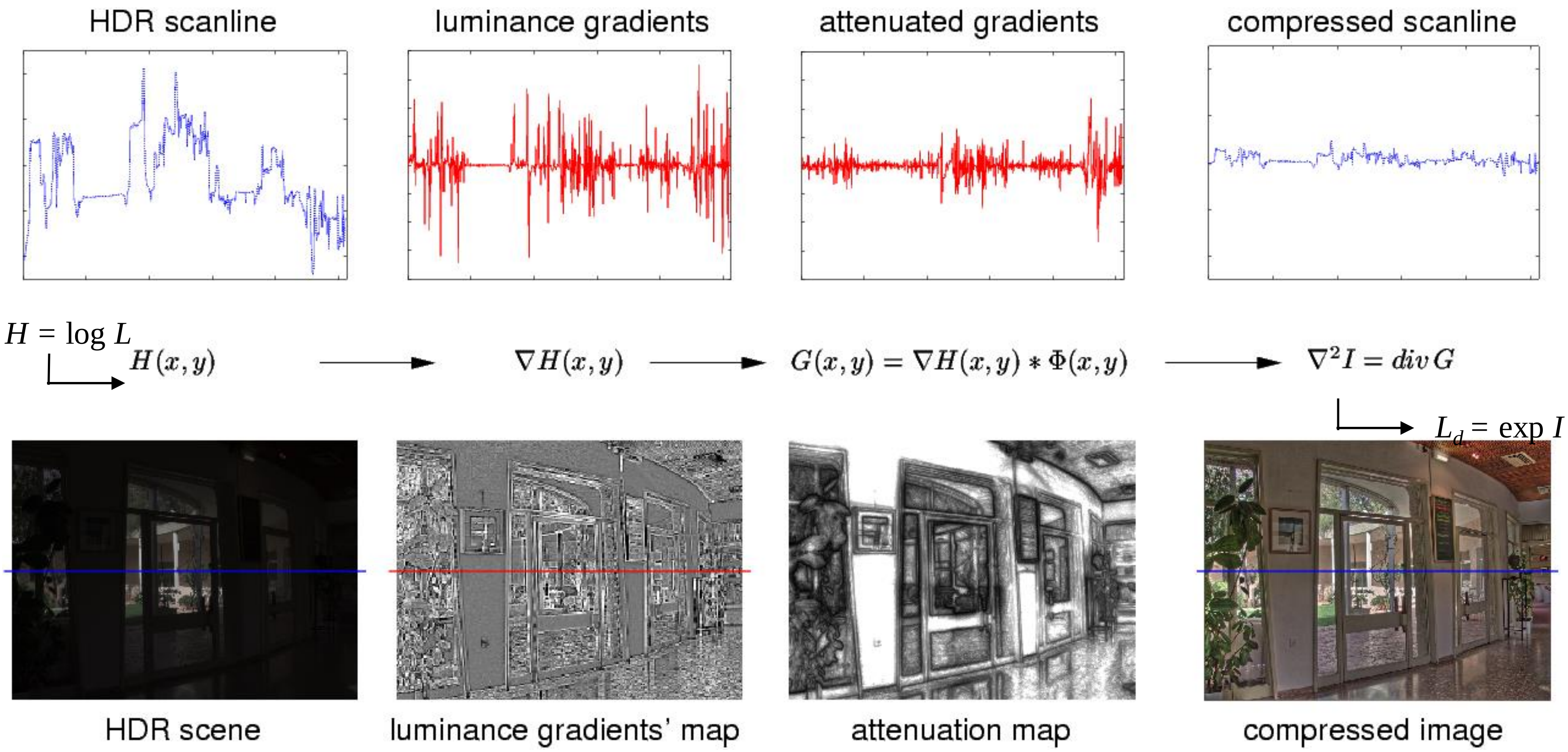
Tone Mapping Algorithm



Luminance in logarithmic domain.

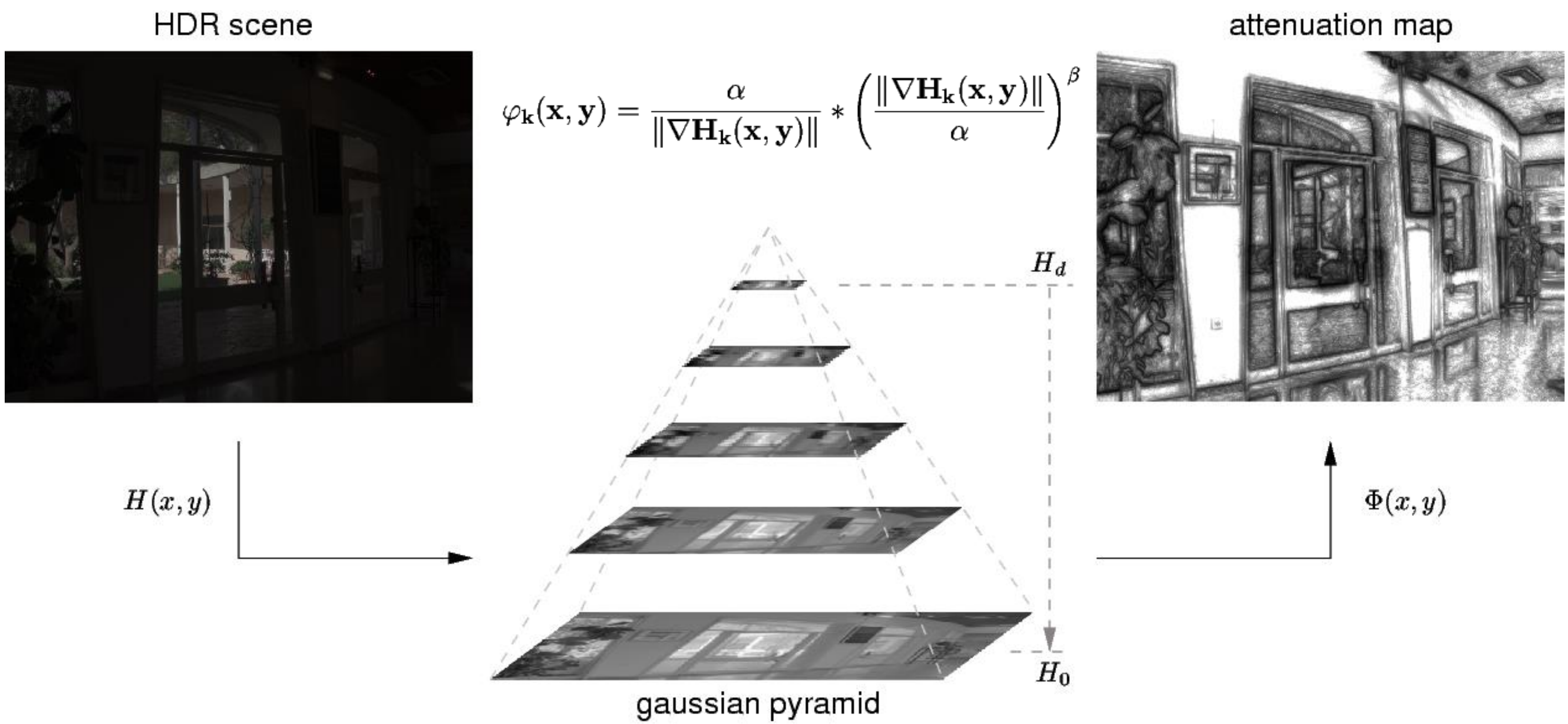
Illumination & Reflectance

Gradient Compression Algorithm



1. Calculate gradients map of image
2. Calculate attenuation map
3. Attenuate gradients
4. Solve Poisson equation to recover image

Attenuation Map

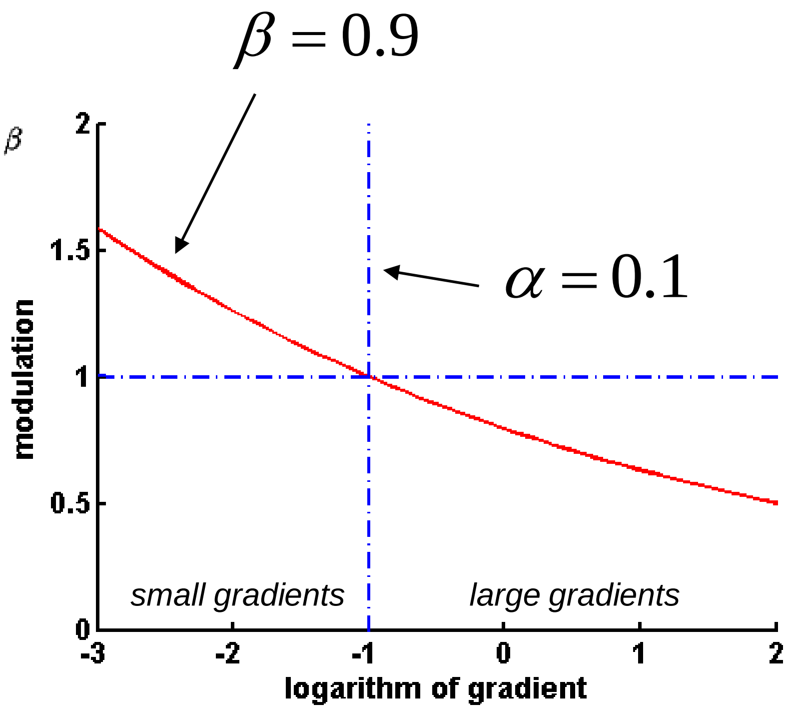


$$\varphi_{\mathbf{k}}(\mathbf{x},\mathbf{y}) = \frac{\alpha}{\|\nabla \mathbf{H_k}(\mathbf{x},\mathbf{y})\|} * \left(\frac{\|\nabla \mathbf{H_k}(\mathbf{x},\mathbf{y})\|}{\alpha}\right)^{\beta}$$

1. Create Gaussian pyramid
2. Calculate gradients on levels
3. Calculate attenuation on levels - φ_k
4. Propagate levels to full resolution

Transfer Function for Contrasts

$$\varphi_{\mathbf{k}}(\mathbf{x},\mathbf{y}) = \frac{\alpha}{\|\nabla \mathbf{H}_{\mathbf{k}}(\mathbf{x},\mathbf{y})\|} * \left(\frac{\|\nabla \mathbf{H}_{\mathbf{k}}(\mathbf{x},\mathbf{y})\|}{\alpha}\right)^{\beta}$$



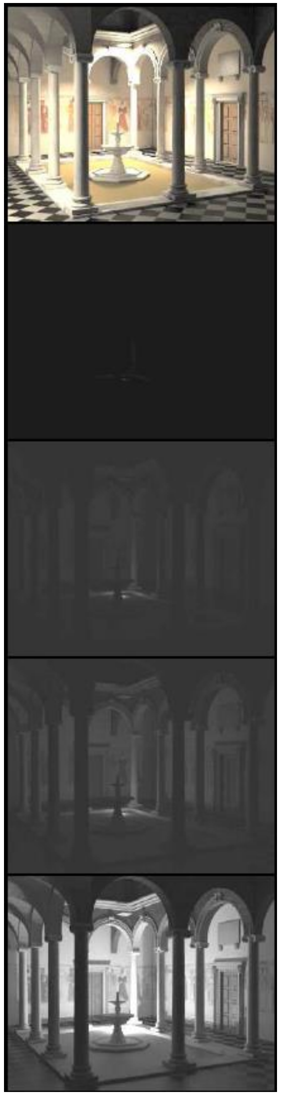
- Attenuate large gradients
  - presumably illumination
- Amplify small gradients
  - hopefully texture details
  - but also noise
- Equation has a division by zero!

Perceptual Effects in TM

- Simulate effects that do not appear on a screen but are typically observed in real-world scenes
  - veiling glare
  - night vision
  - temporal adaptation to light
- Increase believability of results, because we associate such effects with luminance conditions



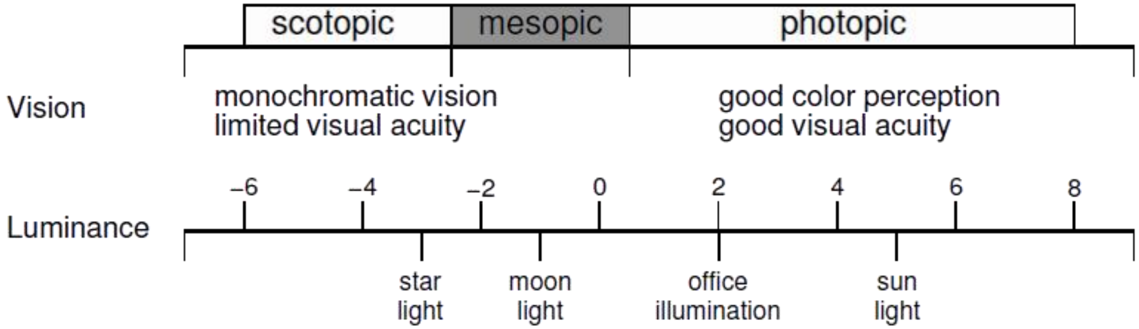
Temporal Luminance Adaptation



- Compensates changes in illumination
- Simulated by smoothing adapting luminance in tone mapping equation
- Different speed of adaptation to light and to darkness

Night Vision

- Human Vision operates in three distinct adaptation conditions:



Visual Acuity

- Perception of spatial details is limited with decreasing illumination level
- Details can be removed using convolution with a Gaussian kernel
- Highest resolvable spatial frequency:

$$RF(Y) = 17.25 \cdot \arctan(1.4 \log_{10} Y + 0.35) + 25.72$$

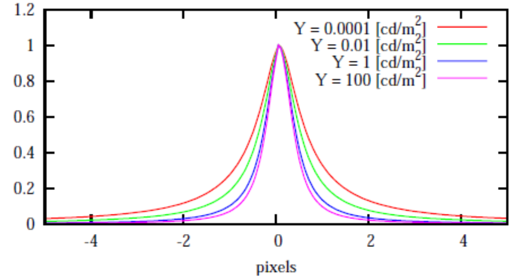
Veiling Luminance (Glare)

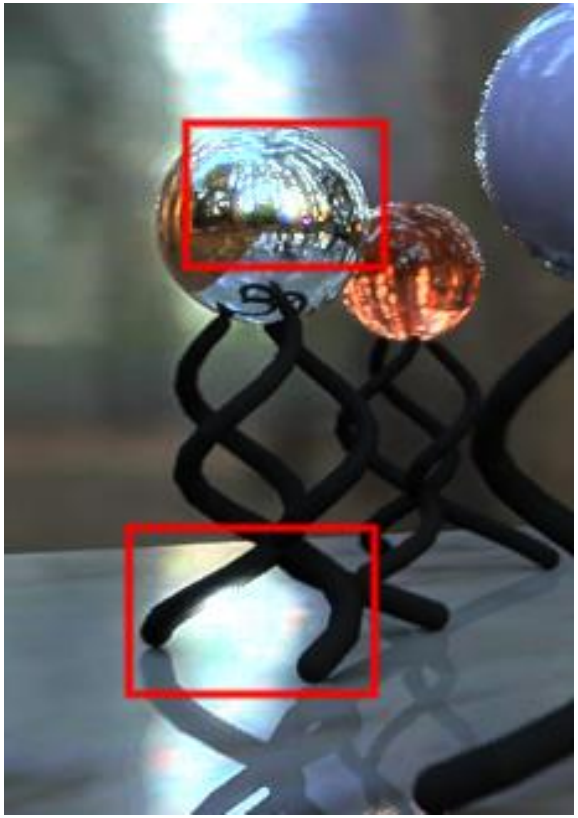
- Decrease of contrast and visibility due to light scattering in the optical system of the eye

- Described by the optical transfer function:

$$OTF(\rho, d(\bar{Y})) = \exp\left(-\frac{\rho}{20.9 - 2.1 \cdot d}^{1.3 - 0.07 \cdot d}\right)$$

ρ spatial frequency, d pupil aperture





HDR Video Player with Perceptual Effects

Papers on Multi-exposure Techniques

- Estimation-Theoretic Approach to Dynamic Range Improvement Using Multiple Exposures
  - M. Robertson, S. Borman, and R. Stevenson
  - In: Journal of Electronic Imaging, vol. 12(2), April 2003.
- Recovering High Dynamic Range Radiance Maps from Photographs
  - Paul E. Debevec and Jitendra Malik
  - In: SIGGRAPH 97
- Radiometric Self Calibration
  - T. Mitsunaga and S.K. Nayar
  - In: Computer Vision and Pattern Recognition (CVPR), 1999.
- High Dynamic Range from Multiple Images: Which Exposures to Combine?
  - M.D. Grossberg and S.K. Nayar
  - In: ICCV Workshop on Color and Photometric Methods in Computer Vision (CPMCV), 2003.

Papers about Tone Mapping

- Adaptive Logarithmic Mapping for Displaying High Contrast Scenes
  - F. Drago, K. Myszkowski, T. Annen, and N. Chiba
  - In: Eurographics 2003
- Photographic Tone Reproduction for Digital Images
  - E. Reinhard, M. Stark, P. Shirley, and J. Ferwerda
  - In: SIGGRAPH 2002 (ACM Transactions on Graphics)
- Fast Bilateral Filtering for the Display of High-Dynamic-Range Images
  - F. Durand and J. Dorsey
  - In: SIGGRAPH 2002 (ACM Transactions on Graphics)
- Gradient Domain High Dynamic Range Compression
  - R. Fattal, D. Lischinski, and M. Werman
  - In: SIGGRAPH 2002 (ACM Transactions on Graphics)
- Dynamic Range Reduction Inspired by Photoreceptor Physiology
  - E. Reinhard and K. Devlin
  - In IEEE Transactions on Visualization and Computer Graphics, 2005
- Time-Dependent Visual Adaptation for Realistic Image Display
  - S.N. Pattanaik, J. Tumblin, H. Yee, and D.P. Greenberg
  - In: Proceedings of ACM SIGGRAPH 2000
- Lightness Perception in Tone Reproduction for High Dynamic Range Images
  - G. Krawczyk, K. Myszkowski, H.-P. Seidel
  - In: Eurographics 2005
- Perceptual Effects in Real-time Tone Mapping
  - G. Krawczyk, K. Myszkowski, H.-P. Seidel
  - In: Spring Conference on Computer Graphics, 2005

Acknowledgements

- I would like to thank Grzesiek Krawczyk for making his slides available.